Basic Concepts in Optimization – Part I

Benoît Chachuat <benoit@mcmaster.ca>

McMaster University
Department of Chemical Engineering

ChE 4G03: Optimization in Chemical Engineering

Outline

Local Optima

Neighborhood

The neighborhood $N_\delta(\mathbf{x}^\circ)$ of a point $\mathbf{x}^\circ$ consists of all nearby points; that is, all points within a small distance $\delta > 0$ of $\mathbf{x}^\circ$:

$$N_\delta(\mathbf{x}^\circ) \triangleq \{\mathbf{x} : \|\mathbf{x} - \mathbf{x}^\circ\| < \delta\}$$

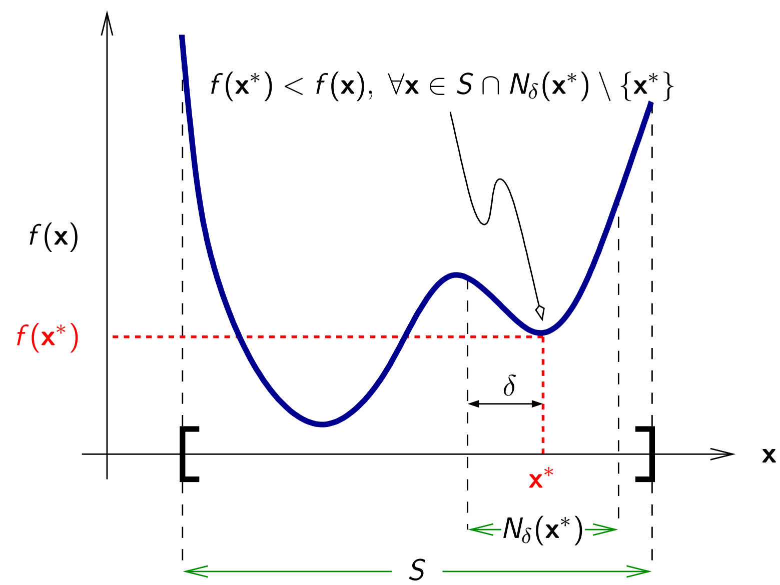
Local Optimum

A point $\mathbf{x}^*$ is a [strict] local minimum for the function $f : \mathbb{R}^n \to \mathbb{R}$ on the set S if it is feasible ($\mathbf{x}^* \in S$) and if sufficiently small neighborhoods surrounding it contain no points that are both feasible and [strictly] lower in objective value:

$$\exists \delta > 0 : \quad f(\mathbf{x}^*) \leq f(\mathbf{x}), \quad \forall \mathbf{x} \in S \cap N_\delta(\mathbf{x}^*)$$

$$[\ \exists \delta > 0 : \quad f(\mathbf{x}^*) < f(\mathbf{x}), \quad \forall \mathbf{x} \in S \cap N_\delta(\mathbf{x}^*) \setminus \{\mathbf{x}^*\}\ ]$$

Illustration of a (Strict) Local Minimum, x^*



Global Optima

Global Optimum

A point $\mathbf{x}^*$ is a [strict] global minimum for the function $f : \mathbb{R}^n \to \mathbb{R}$ on the set S if it is feasible ($\mathbf{x}^* \in S$) and if no other feasible solution has [strictly] lower objective value:

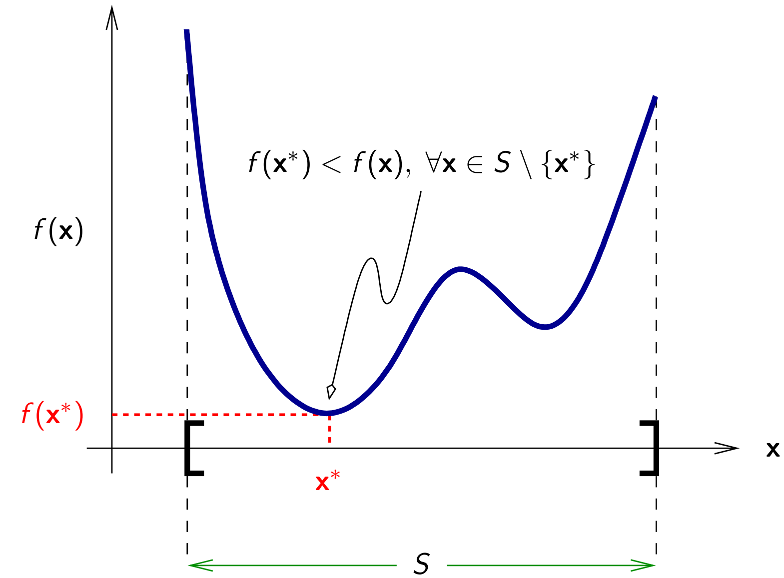
$$f(\mathbf{x}^*) \leq f(\mathbf{x}), \quad \forall \mathbf{x} \in S$$

$$[\, f(\mathbf{x}^*) < f(\mathbf{x}), \quad \forall \mathbf{x} \in S \setminus \{\mathbf{x}^*\} \,]$$

Remarks:

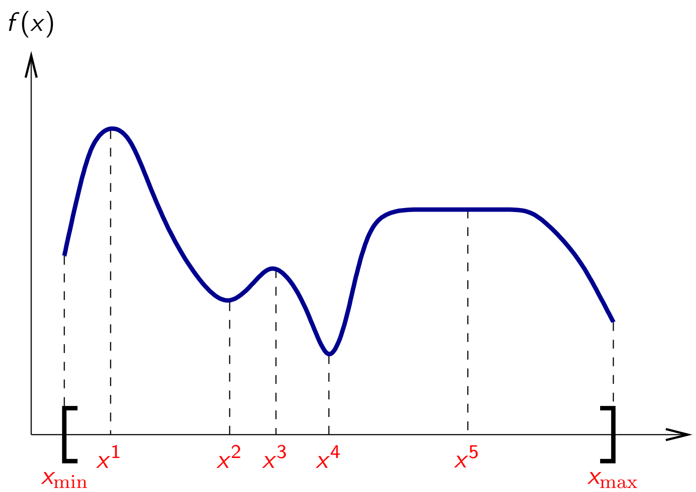
1. Global minima are always local minima
2. Local minima may not be global minima
3. Analog definitions hold for local/global optima to maximize problems

Illustration of a (Strict) Global Minimum, x^*



Global vs. Local Optima

Class Exercise: Identify the various types of minima and maxima for f on $S \triangleq [x_{\min}, x_{\max}]$



How to Find Optima?

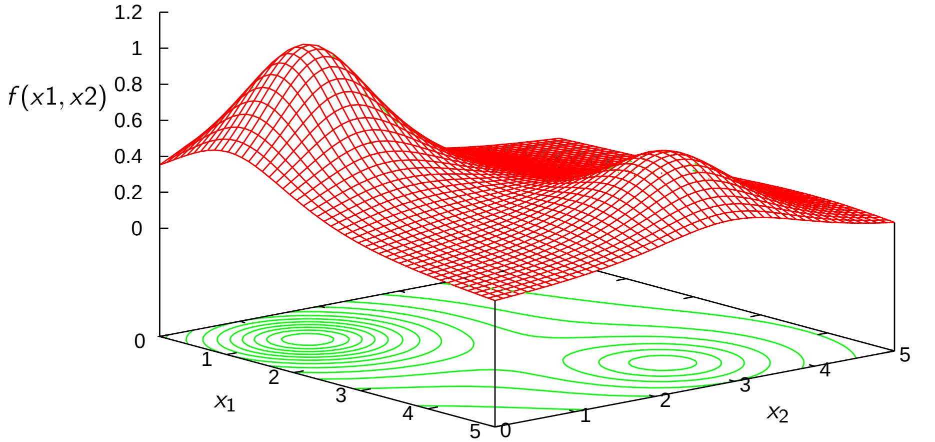
Review: Three Methods for Optimization

1. **Graphical Solutions**
   Great display + see multiple optima
   But impractical for nearly all practical problems
2. **Analytical Solutions** (e.g., Newton, Euler, etc.)
   Exact solution + easy analysis for changes in (uncertain) parameters
   But not possible for most practical problems
3. **Numerical Solutions**
   The only practical method for complex models!
   But only guarantees local optima + challenges in finding effects of (uncertain) parameters

Numerical Optimization: The Dilemma!

Consider the optimization problem:

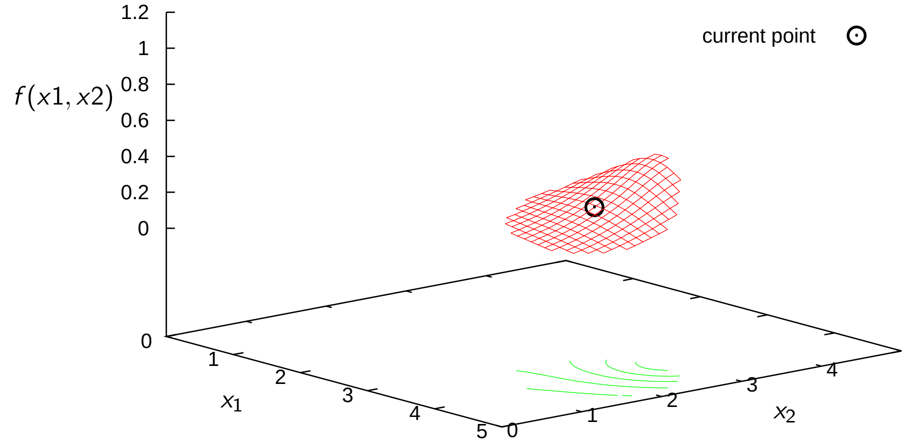
$$\min_{0 \leq x_1, x_2 \leq 5} f(x_1, x_2) \triangleq \frac{1}{1 + (x_1 - 1)^2 + (x_2 - 1)^2} + \frac{0.5}{1 + (x_1 - 4)^2 + (x_2 - 3)^2}$$



Numerical Optimization: The Dilemma!

Typically, only some local information is know about the objective function typically at a current point $\mathbf{x}^\circ = (x_1^\circ, x_2^\circ)$!

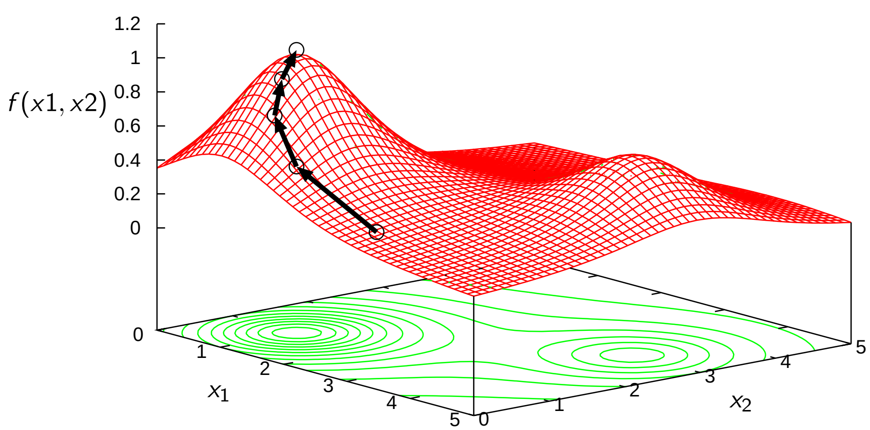
Question: Which move do I make next?



Numerical Optimization: The Basic Approach

Improving Search

Improving search methods are numerical algorithms that begin at a feasible solution to a given optimization model, and advance along a search path of feasible points with ever-improving function value



Direction-Step Paradigm

At the current point $\mathbf{x}^{(k)}$, how do I decide:

- the direction of change
- the magnitude of change
- whether further improvement is possible?

The Basic Equation

Improving search advances from current point $\mathbf{x}^{(k)}$ to new point $\mathbf{x}^{(k+1)}$ as:

$$\mathbf{x}^{(k+1)} = \begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ \vdots \\ x_n^{(k+1)} \end{pmatrix} = \mathbf{x}^{(k)} + \alpha \boldsymbol{\Delta} \mathbf{x} = \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \vdots \\ x_n^{(k)} \end{pmatrix} + \alpha \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_n \end{pmatrix}$$

where:

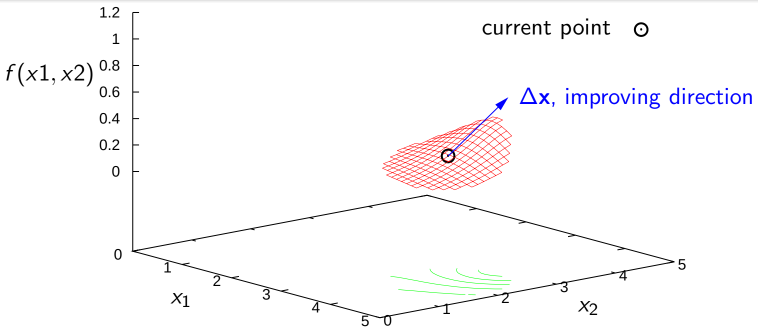
- $\Delta \mathbf{x}$ defines a move direction of solution change at $\mathbf{x}^{(k)}$ ($\|\Delta \mathbf{x}\| = 1$)
- $\alpha > 0$ determines a move magnitude, how far to pursue this direction

Direction of Change, $\Delta\mathbf{x}$

Improving Directions

Vector $\Delta\mathbf{x} \in \mathbb{R}^n$ is an improving direction at current point $\mathbf{x}^{(k)}$ if the objective function value at $\mathbf{x}^{(k)} + \alpha\Delta\mathbf{x}$ is superior to that of $\mathbf{x}^{(k)}$, for *all* $\alpha > 0$ sufficiently small

$$\text{(maximize problem)} \quad \exists\bar{\alpha} > 0: \quad f(\mathbf{x}^{(k)} + \alpha\Delta\mathbf{x}) > f(\mathbf{x}^{(k)}), \quad \forall\alpha \in (0, \bar{\alpha}]$$

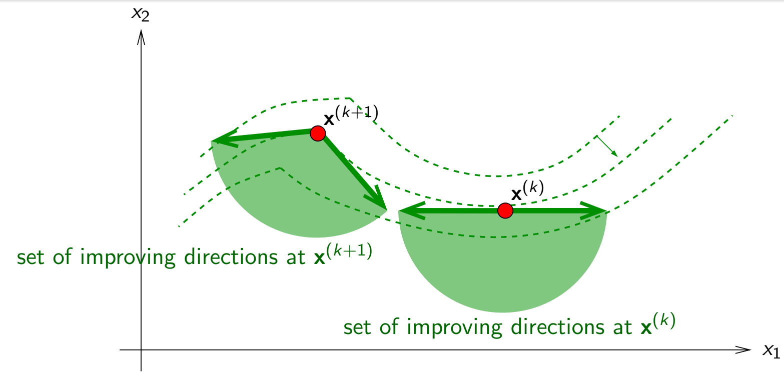


Direction of Change, $\Delta\mathbf{x}$

Improving Directions

Vector $\Delta\mathbf{x} \in \mathbb{R}^n$ is an improving direction at current point $\mathbf{x}^{(k)}$ if the objective function value at $\mathbf{x}^{(k)} + \alpha\Delta\mathbf{x}$ is superior to that of $\mathbf{x}^{(k)}$, for *all* $\alpha > 0$ sufficiently small

$$\text{(maximize problem)} \quad \exists\bar{\alpha} > 0: \quad f(\mathbf{x}^{(k)} + \alpha\Delta\mathbf{x}) > f(\mathbf{x}^{(k)}), \quad \forall\alpha \in (0, \bar{\alpha}]$$

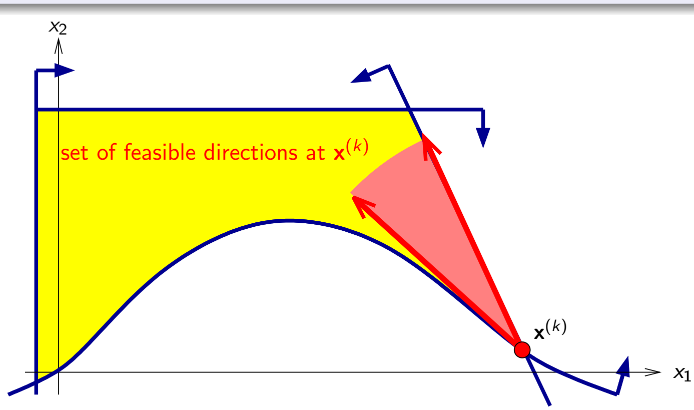


Direction of Change, $\Delta\mathbf{x}$ (cont'd)

Feasible Directions

Vector $\Delta\mathbf{x} \in \mathbb{R}^n$ is an feasible direction at current point $\mathbf{x}^{(k)}$ if point $\mathbf{x}^{(k)} + \alpha\Delta\mathbf{x}$ violates no model constraint for *all* $\alpha > 0$ sufficiently small

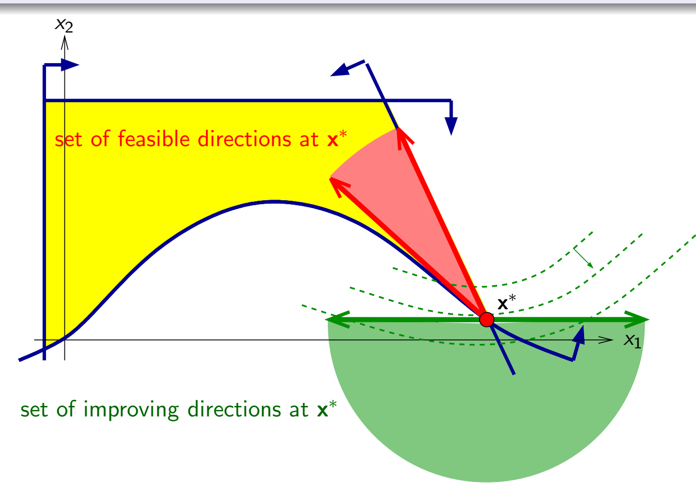
$$\exists\bar{\alpha} > 0: \quad \mathbf{x}^{(k)} + \alpha\Delta\mathbf{x} \in S, \quad \forall\alpha \in (0, \bar{\alpha}]$$



Optimality Criterion

Necessary Condition of Optimality (NCO)

No optimization model solution at which an improving feasible direction is available can be a local optimum



Continuous Improving Search Algorithm

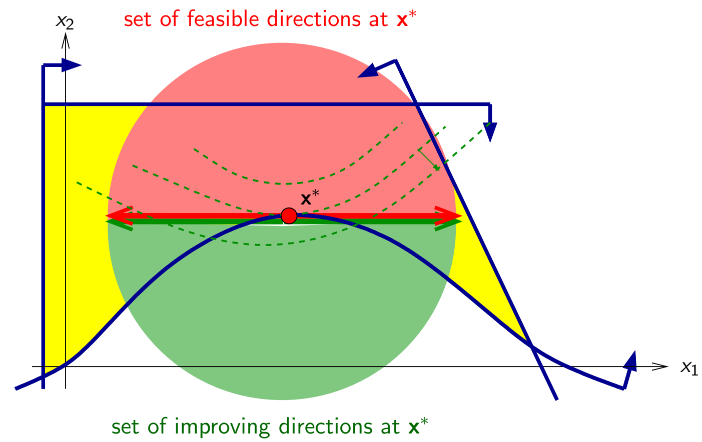
- **Step 0: Initialization.**
  - Choose any starting feasible point $\mathbf{x}^{(0)}$ and let index $k \leftarrow 0$.
- **Step 1: Move Direction.**
  - If no improving feasible direction $\Delta\mathbf{x}$ exists at current point $\mathbf{x}^{(k)}$, stop.
  - Otherwise, construct an improving feasible direction at $\mathbf{x}^{(k)}$ as $\Delta\mathbf{x}^{(k+1)}$.
- **Step 2: Step Size.**
  - If there is no limit on step sizes for which direction $\Delta\mathbf{x}^{(k+1)}$ continues to both improve the objective function and retain feasibility, stop — The model is unbounded.
  - Otherwise, choose the largest step size $\alpha^{(k+1)}$.
- **Step 3: Update.**
  - $\mathbf{x}^{(k+1)} \leftarrow \mathbf{x}^{(k)} + \alpha^{(k+1)}\Delta\mathbf{x}^{(k+1)}$
  - Increment index $k \leftarrow k+1$ and return to step 1.

Remarks:

- This basic algorithm may terminate at a suboptimal point
- Moreover, it does not distinguish between local and global optima

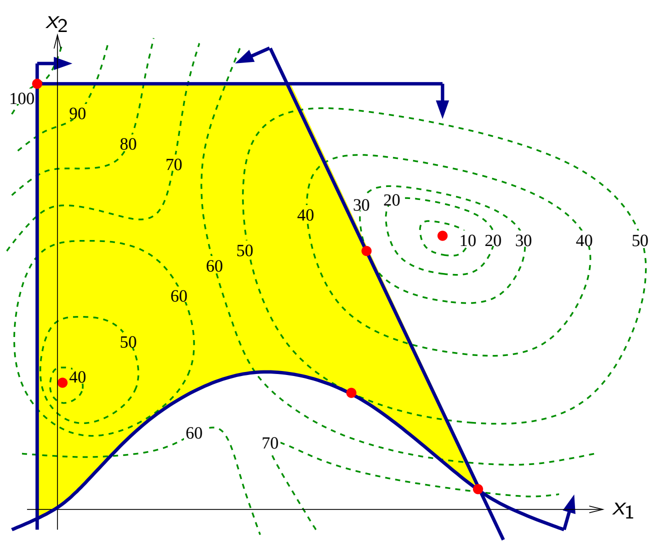
A Word of Caution!

Caution: A point at which no improving feasible direction is available may **not** be a local optimum!



Finding out Optima!

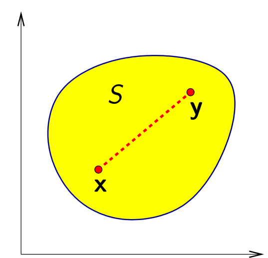
Class Exercise: Determine whether each of the following points is apparently a local/global minimum? a local/global maximum? neither?



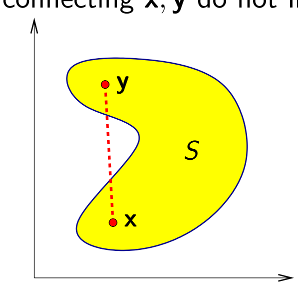
Convex Sets

A set $S \subset \mathbb{R}^n$ is said to be **convex** if every point on the line connecting any two points $\mathbf{x}, \mathbf{y}$ in S is itself in S,

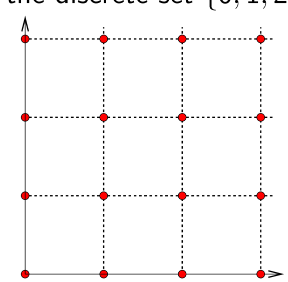
$$\gamma\mathbf{x} + (1-\gamma)\mathbf{y} \in S, \quad \forall \gamma \in (0,1)$$



Nonconvex Set: Some points on the line connecting $\mathbf{x}, \mathbf{y}$ do not lie in S



Nonconnected sets are nonconvex; e.g., the discrete set $\{0, 1, 2, \ldots\}^2$



Convex and Concave Functions

Convex Functions

A function $f : S \to \mathbb{R}$, defined on a convex set $S \subset \mathbb{R}^n$, is said to be **convex on** S if the line segment connecting $f(\mathbf{x})$ and $f(\mathbf{y})$ at **any** two points $\mathbf{x}, \mathbf{y} \in S$ lies above the function between $\mathbf{x}$ and $\mathbf{y}$,

$$f(\gamma\mathbf{x} + (1-\gamma)\mathbf{y}) \leq \gamma f(\mathbf{x}) + (1-\gamma) f(\mathbf{y}), \quad \forall \gamma \in (0,1)$$

- **Strict** convexity:

$$f(\gamma\mathbf{x} + (1-\gamma)\mathbf{y}) < \gamma f(\mathbf{x}) + (1-\gamma) f(\mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y} \in S, \quad \forall \gamma \in (0,1)$$

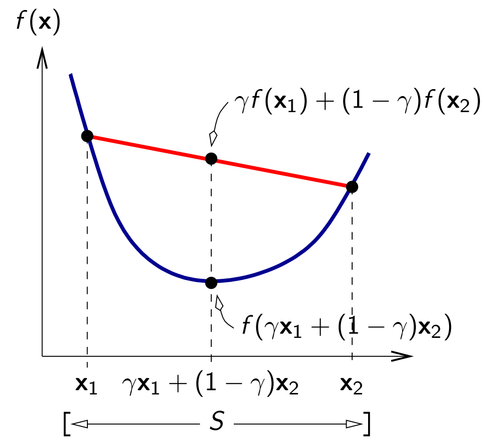
Concave Functions

f is said to be **[strictly] concave on** S if $(-f)$ is [strictly] convex on S,

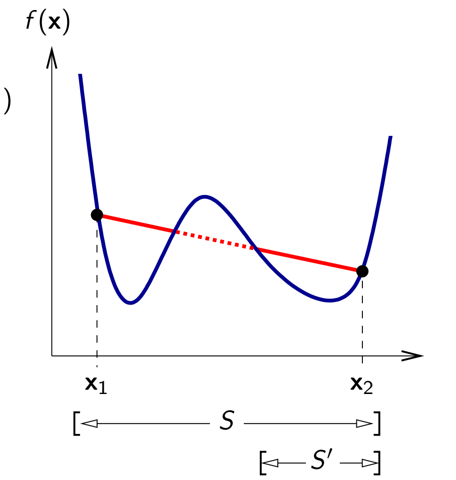
$$f(\gamma\mathbf{x} + (1-\gamma)\mathbf{y}) \geq [>] \gamma f(\mathbf{x}) + (1-\gamma) f(\mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y} \in S, \quad \forall \gamma \in (0,1)$$

Convex and Concave Functions (cont'd)

Case of a strictly convex function on the convex set S

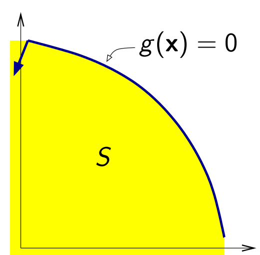


Case of a nonconvex function on S, yet convex on the convex set S'



Sets Defined by Constraints

Define the set $S \triangleq \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) \leq 0\}$, with g a convex function on $\mathbb{R}^n$. Then, S is a convex set



Why?

- Consider any two points $\mathbf{x}, \mathbf{y} \in S$. By the convexity of g,
$$g(\gamma\mathbf{x} + (1-\gamma)\mathbf{y}) \leq \gamma g(\mathbf{x}) + (1-\gamma) g(\mathbf{y}), \quad \forall \gamma \in (0,1)$$
- Since $g(\mathbf{x}) \leq 0$ and $g(\mathbf{y}) \leq 0$,
$$g(\mathbf{x}) + (1-\gamma) g(\mathbf{y}) \leq 0, \quad \forall \gamma \in (0,1)$$
- Therefore, $\gamma\mathbf{x} + (1-\gamma)\mathbf{y} \in S$ for every $\gamma \in (0,1)$; i.e., S is convex

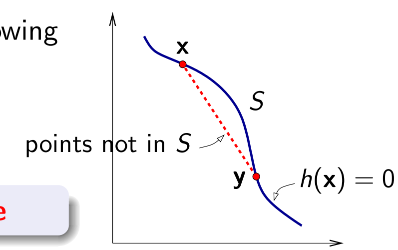
Class Exercise: Give a condition on g for the following set to be convex:

$$S \triangleq \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) \geq 0\}$$

Sets Defined by Constraints (cont'd)

- What is the condition on h for the following set to be convex:

$$S \triangleq \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) = 0\}$$



The set S is convex **if and only if** h **is affine**

Convex Sets Defined by Constraints

Consider the set

$$S \triangleq \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \leq 0, \ldots, g_m(\mathbf{x}) \leq 0, h_1(\mathbf{x}) = 0, \ldots, h_p(\mathbf{x}) = 0\}$$

Then, S is convex if:

- $g_1, \ldots, g_m$ are convex on $\mathbb{R}^n$
- $h_1, \ldots, h_p$ are affine

Convexity and Global Optimality

- Consider the constrained program:

$$\begin{aligned}\max_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & g_j(\mathbf{x}) \leq 0, \quad j = 1, \ldots, m \\ & h_j(\mathbf{x}) = 0, \quad j = 1, \ldots, p\end{aligned}$$

- If f and $g_1, \ldots, g_m$ are convex on $\mathbb{R}^n$, and $h_1, \ldots, h_p$ are affine, then this program is said to be a **convex program**

Sufficient Condition for **Global** Optimality

A [strict] local minimum to a convex program is also a [strict] global minimum

- On the other hand, a nonconvex program may or may not have local optima that are not global optima